

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 10/783,206
Source: EFW
Date Processed by STIC: 11-17-04

ENTERED



IFWO

RAW SEQUENCE LISTING

DATE: 11/17/2004

PATENT APPLICATION: US/10/783,206

TIME: 12:20:52

Input Set : A:\Ipt01202.app

Output Set: N:\CRF4\11172004\J783206.raw

3 <110> APPLICANT: DARST, SETH A.
 4 CAMPBELL, ELIZABETH
 6 <120> TITLE OF INVENTION: A CRYSTAL OF BACTERIAL CORE RNA POLYMERASE WITH
 7 RIFAMPICIN AND METHODS OF USE THEREOF
 9 <130> FILE REFERENCE: IPT-012.02
 11 <140> CURRENT APPLICATION NUMBER: 10/783,206
 12 <141> CURRENT FILING DATE: 2004-02-20
 14 <150> PRIOR APPLICATION NUMBER: 09/802,755
 15 <151> PRIOR FILING DATE: 2001-03-09
 17 <160> NUMBER OF SEQ ID NOS: 30
 19 <170> SOFTWARE: PatentIn Ver. 3.2
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 1525
 23 <212> TYPE: PRT
 24 <213> ORGANISM: Thermus aquaticus
 26 <220> FEATURE:
 27 <221> NAME/KEY: MOD_RES
 28 <222> LOCATION: (1247)
 29 <223> OTHER INFORMATION: Variable amino acid
 31 <400> SEQUENCE: 1
 32 Met Lys Lys Glu Val Arg Lys Val Arg Ile Ala Leu Ala Ser Pro Glu
 33 1 5 10 15
 35 Lys Ile Arg Ser Trp Ser Tyr Gly Glu Val Glu Lys Pro Glu Thr Ile
 36 20 25 30
 38 Asn Tyr Arg Thr Leu Lys Pro Glu Arg Asp Gly Leu Phe Asp Glu Arg
 39 35 40 45
 41 Ile Phe Gly Pro Ile Lys Asp Tyr Glu Cys Ala Cys Gly Lys Tyr Lys
 42 50 55 60
 44 Arg Gln Arg Phe Glu Gly Lys Val Cys Glu Arg Cys Gly Val Glu Val
 45 65 70 75 80
 47 Thr Arg Ser Ile Val Arg Arg Tyr Arg Met Gly His Ile Glu Leu Ala
 48 85 90 95
 50 Thr Pro Ala Ala His Ile Trp Phe Val Lys Asp Val Pro Ser Lys Ile
 51 100 105 110
 53 Gly Thr Leu Leu Asp Leu Phe Ala Thr Glu Leu Glu Gln Val Leu Tyr
 54 115 120 125
 56 Phe Asn Lys Tyr Ile Val Leu Asp Pro Lys Gly Ala Val Leu Asp Gly
 57 130 135 140
 59 Val Pro Val Glu Lys Arg Gln Leu Leu Thr Asp Glu Glu Tyr Arg Glu
 60 145 150 155 160
 62 Leu Arg Tyr Gly Lys Gln Glu Thr Tyr Pro Leu Pro Ala Gly Val Asp
 63 165 170 175
 65 Ala Leu Val Lys Asp Gly Glu Glu Val Val Lys Gly Gln Glu Leu Ala

(ps. 6)

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/783,206

DATE: 11/17/2004

TIME: 12:20:52

Input Set : A:\Ipt01202.app

Output Set: N:\CRF4\11172004\J783206.raw

```

66          180          185          190
68 Pro Gly Val Val Ser Arg Met Asp Gly Val Gly Ser Leu Pro Leu Pro
69          195          200          205
71 Arg Arg Val Arg Val Asp Tyr Leu Arg Lys Glu Arg Ala Ala Leu Arg
72          210          215          220
74 Ile Pro Leu Ser Ala Trp Val Glu Lys Glu Pro Tyr Arg Pro Gly Glu
75 225          230          235          240
77 Val Leu Ala Glu Leu Ser Glu Pro Tyr Leu Phe Arg Ala Glu Glu Ser
78          245          250          255
80 Gly Val Val Glu Leu Lys Asp Leu Ala Glu Gly His Leu Ile Tyr Leu
81          260          265          270
83 Arg Gln Glu Glu Glu Val Val Ala Arg Tyr Phe Leu Pro Ala Gly Met
84          275          280          285
86 Thr Pro Leu Val Val Glu Gly Glu Ile Val Glu Val Gly Gln Pro Leu
87          290          295          300
89 Ala Glu Gly Lys Gly Leu Leu Arg Leu Pro Arg His Met Thr Ala Lys
90 305          310          315          320
92 Glu Val Glu Ala Glu Glu Glu Gly Asp Ser Val His Leu Thr Leu Phe
93          325          330          335
95 Leu Glu Trp Thr Glu Pro Lys Asp Tyr Lys Val Ala Pro His Met Asn
96          340          345          350
98 Val Ile Val Pro Glu Gly Ala Lys Val Gln Ala Gly Glu Lys Ile Val
99          355          360          365
101 Ala Ala Ile Asp Pro Glu Glu Val Ile Ala Gln Ala Glu Gly Val
102          370          375          380
104 Val His Leu His Glu Pro Ala Ser Ile Leu Val Val Lys Ala Arg Val
105 385          390          395          400
107 Tyr Pro Phe Glu Asp Asp Val Glu Val Thr Thr Gly Asp Arg Val Ala
108          405          410          415
110 Pro Gly Asp Val Leu Ala Asp Gly Gly Lys Val Lys Ser Glu Ile Tyr
111          420          425          430
113 Gly Arg Val Glu Val Asp Leu Val Arg Asn Val Val Arg Val Val Glu
114          435          440          445
116 Ser Tyr Asp Ile Asp Ala Arg Met Gly Ala Glu Ala Ile Gln Glu Leu
117          450          455          460
119 Leu Lys Glu Leu Asp Leu Glu Lys Leu Glu Arg Glu Leu Leu Glu Glu
120 465          470          475          480
122 Met Lys His Pro Ser Arg Ala Arg Arg Ala Lys Ala Arg Lys Arg Leu
123          485          490          495
125 Glu Val Val Arg Ala Phe Leu Asp Ser Gly Asn Arg Pro Glu Trp Met
126          500          505          510
128 Ile Leu Glu Ala Val Pro Val Leu Pro Pro Asp Leu Arg Pro Met Val
129          515          520          525
131 Gln Val Asp Gly Gly Arg Phe Ala Thr Ser Asp Leu Asn Asp Leu Tyr
132          530          535          540
134 Arg Arg Leu Ile Asn Arg Asn Asn Arg Leu Lys Lys Leu Leu Ala Gln
135 545          550          555          560
137 Gly Ala Pro Glu Ile Ile Ile Arg Asn Glu Lys Arg Met Leu Gln Glu
138          565          570          575

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/783,206

DATE: 11/17/2004

TIME: 12:20:52

Input Set : A:\Ipt01202.app

Output Set: N:\CRF4\11172004\J783206.raw

```

140 Ala Val Asp Ala Val Ile Asp Asn Gly Arg Arg Gly Ser Pro Val Thr
141          580          585          590
143 Asn Pro Gly Ser Glu Arg Pro Leu Arg Ser Leu Thr Asp Ile Leu Ser
144          595          600          605
146 Gly Lys Gln Gly Arg Phe Arg Gln Asn Leu Leu Gly Lys Arg Val Asp
147          610          615          620
149 Tyr Ser Gly Arg Ser Val Ile Val Val Gly Pro Gln Leu Lys Leu His
150 625          630          635          640
152 Gln Cys Gly Leu Pro Lys Arg Met Ala Leu Glu Leu Phe Lys Pro Phe
153          645          650          655
155 Leu Leu Lys Lys Met Glu Glu Lys Ala Phe Ala Pro Asn Val Lys Ala
156          660          665          670
158 Ala Arg Arg Met Leu Glu Arg Gln Arg Asp Ile Lys Asp Glu Val Trp
159          675          680          685
161 Asp Ala Leu Glu Glu Val Ile His Gly Lys Val Val Leu Leu Asn Arg
162          690          695          700
164 Ala Pro Thr Leu His Arg Leu Gly Ile Gln Ala Phe Gln Pro Val Leu
165 705          710          715          720
167 Val Glu Gly Gln Ser Ile Gln Leu His Pro Leu Val Cys Glu Ala Phe
168          725          730          735
170 Asn Ala Asp Phe Asp Gly Asp Gln Met Ala Val His Val Pro Leu Ser
171          740          745          750
173 Ser Phe Ala Gln Ala Glu Ala Arg Ile Gln Met Leu Ser Ala His Asn
174          755          760          765
176 Leu Leu Ser Pro Ala Ser Gly Glu Pro Leu Ala Lys Pro Ser Arg Asp
177          770          775          780
179 Ile Ile Leu Gly Leu Tyr Tyr Ile Thr Gln Val Arg Lys Glu Lys Lys
180 785          790          795          800
182 Gly Ala Gly Met Ala Phe Ala Thr Pro Glu Glu Ala Leu Ala Ala Tyr
183          805          810          815
185 Glu Arg Gly Glu Val Ala Leu Asn Ala Pro Ile Val Val Ala Gly Arg
186          820          825          830
188 Glu Thr Ser Val Gly Arg Leu Lys Phe Val Phe Ala Asn Pro Asp Glu
189          835          840          845
191 Ala Leu Leu Ala Val Ala His Gly Leu Leu Asp Leu Gln Asp Val Val
192          850          855          860
194 Thr Val Arg Tyr Leu Gly Arg Arg Leu Glu Thr Asn Pro Gly Arg Ile
195 865          870          875          880
197 Leu Phe Ala Arg Ile Val Gly Glu Ala Val Gly Asp Glu Lys Val Ala
198          885          890          895
200 Gln Glu Leu Ile Gln Met Asp Val Pro Gln Glu Lys Asn Ser Leu Lys
201          900          905          910
203 Asp Leu Val Tyr Gln Ala Phe Leu Arg Leu Gly Met Glu Lys Thr Ala
204          915          920          925
206 Arg Leu Leu Asp Ala Leu Lys Tyr Tyr Gly Phe Thr Leu Ser Thr Thr
207          930          935          940
209 Ser Gly Ile Ile Thr Ile Gly Ile Asp Asp Ala Val Ile Pro Glu Glu
210 945          950          955          960
212 Lys Gln Arg Tyr Leu Glu Glu Ala Asp Arg Lys Leu Arg Gln Ile Glu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/783,206

DATE: 11/17/2004

TIME: 12:20:52

Input Set : A:\Ipt01202.app

Output Set: N:\CRF4\11172004\J783206.raw

```

213          965          970          975
215 Gln Ala Tyr Glu Met Gly Phe Leu Thr Asp Arg Glu Arg Tyr Asp Gln
216          980          985          990
218 Val Ile Gln Leu Trp Thr Glu Thr Thr Glu Lys Val Thr Gln Ala Val
219          995          1000          1005
221 Phe Asn Asn Phe Glu Glu Asn Tyr Pro Phe Asn Pro Leu Tyr Val Met
222          1010          1015          1020
224 Ala Gln Ser Gly Ala Arg Gly Asn Pro Gln Gln Ile Arg Gln Leu Cys
225          1025          1030          1035          1040
227 Gly Met Arg Gly Leu Met Gln Lys Pro Ser Gly Glu Thr Phe Glu Val
228          1045          1050          1055
230 Pro Val Arg Ser Ser Phe Arg Glu Gly Leu Thr Val Leu Glu Tyr Phe
231          1060          1065          1070
233 Ile Ser Ser His Gly Ala Arg Lys Gly Gly Ala Asp Thr Ala Leu Arg
234          1075          1080          1085
236 Thr Ala Asp Ser Gly Tyr Leu Thr Arg Lys Leu Val Asp Val Ala His
237          1090          1095          1100
239 Glu Ile Val Val Arg Glu Ala Asp Cys Gly Thr Thr Lys Tyr Ile Ser
240          1105          1110          1115          1120
242 Val Pro Leu Phe Gln Met Asp Glu Val Thr Arg Thr Leu Arg Leu Arg
243          1125          1130          1135
245 Lys Arg Ser Asp Ile Glu Ser Gly Leu Tyr Gly Arg Val Leu Ala Arg
246          1140          1145          1150
248 Glu Val Glu Ala Leu Gly Arg Arg Leu Glu Glu Gly Arg Tyr Leu Ser
249          1155          1160          1165
251 Leu Glu Asp Val His Phe Leu Ile Lys Ala Ala Glu Ala Gly Glu Val
252          1170          1175          1180
254 Arg Glu Val Pro Val Arg Ser Pro Leu Thr Cys Gln Thr Arg Tyr Gly
255          1185          1190          1195          1200
257 Val Cys Gln Lys Cys Tyr Gly Tyr Asp Leu Ser Met Ala Arg Pro Val
258          1205          1210          1215
260 Ser Ile Gly Glu Ala Val Gly Val Val Ala Ala Glu Ser Ile Gly Glu
261          1220          1225          1230
W--> 263 Pro Gly Thr Gln Leu Thr Met Arg Thr Phe His Thr Gly Gly Xaa Ala
264          1235          1240          1245
266 Val Gly Thr Asp Ile Thr Gln Gly Leu Pro Arg Val Ile Glu Leu Phe
267          1250          1255          1260
269 Glu Ala Arg Arg Pro Lys Ala Lys Ala Val Ile Ser Glu Ile Asp Gly
270          1265          1270          1275          1280
272 Val Val Arg Ile Glu Gly Glu Asp Arg Leu Ser Val Phe Val Glu
273          1285          1290          1295
275 Ser Glu Gly Phe Ser Lys Glu Tyr Lys Leu Pro Lys Asp Ala Arg Leu
276          1300          1305          1310
278 Leu Val Lys Asp Gly Asp Tyr Val Glu Ala Gly Gln Pro Leu Thr Arg
279          1315          1320          1325
281 Gly Ala Ile Asp Pro His Gln Leu Leu Glu Ala Lys Gly Pro Glu Ala
282          1330          1335          1340
284 Val Glu Arg Tyr Leu Val Asp Glu Ile Gln Lys Val Tyr Arg Ala Gln
285          1345          1350          1355          1360

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/783,206

DATE: 11/17/2004

TIME: 12:20:52

Input Set : A:\Ipt01202.app

Output Set: N:\CRF4\11172004\J783206.raw

```

287 Gly Val Lys Leu His Asp Lys His Ile Glu Ile Val Val Arg Gln Met
288           1365           1370           1375
290 Leu Lys Tyr Val Glu Val Thr Asp Pro Gly Asp Ser Pro Leu Leu Glu
291           1380           1385           1390
293 Gly Gln Val Leu Glu Lys Trp Asp Val Glu Ala Leu Asn Glu Arg Leu
294           1395           1400           1405
296 Ile Ala Glu Gly Lys Val Pro Val Ala Trp Lys Pro Leu Leu Met Gly
297       1410           1415           1420
299 Val Thr Lys Ser Ala Leu Ser Thr Lys Ser Trp Leu Ser Ala Ala Ser
300       1425           1430           1435           1440
302 Phe Gln Asn Thr Thr His Val Leu Thr Glu Ala Ala Ile Ala Gly Lys
303           1445           1450           1455
305 Lys Asp Glu Leu Ile Gly Leu Lys Glu Asn Val Ile Leu Gly Arg Leu
306           1460           1465           1470
308 Ile Pro Ala Gly Thr Gly Ser Asp Phe Val Arg Phe Thr Gln Val Val
309       1475           1480           1485
311 Asp Gln Arg Thr Leu Lys Ala Ile Glu Glu Ala Arg Lys Glu Ala Val
312       1490           1495           1500
314 Glu Ala Lys Glu Lys Glu Ala Pro Arg Arg Pro Val Arg Arg Glu Gln
315       1505           1510           1515           1520
317 Pro Gly Lys Gly Leu
318           1525
321 <210> SEQ ID NO: 2
322 <211> LENGTH: 1119
323 <212> TYPE: PRT
324 <213> ORGANISM: Thermus aquaticus
326 <220> FEATURE:
327 <221> NAME/KEY: MOD_RES
328 <222> LOCATION: (695)..(696)
329 <223> OTHER INFORMATION: Variable amino acid
331 <400> SEQUENCE: 2
332 Met Lys Ile Lys Arg Phe Gly Arg Ile Arg Glu Val Ile Pro Leu Pro
333   1           5           10           15
335 Pro Leu Thr Glu Ile Gln Val Glu Ser Tyr Lys Lys Ala Leu Gln Ala
336           20           25           30
338 Asp Val Pro Pro Glu Lys Arg Glu Asn Val Gly Ile Gln Ala Ala Phe
339           35           40           45
341 Lys Glu Thr Phe Pro Ile Glu Gly Asp Lys Gly Lys Gly Gly Leu
342           50           55           60
344 Val Leu Asp Phe Leu Glu Tyr Arg Ile Gly Asp Pro Pro Phe Ser Gln
345       65           70           75           80
347 Asp Glu Cys Arg Glu Lys Asp Leu Thr Tyr Gln Ala Pro Leu Tyr Ala
348           85           90           95
350 Arg Leu Gln Leu Ile His Lys Asp Thr Gly Leu Ile Lys Glu Asp Glu
351           100           105           110
353 Val Phe Leu Gly His Leu Pro Leu Met Thr Glu Asp Gly Ser Phe Ile
354           115           120           125
356 Ile Asn Gly Ala Asp Arg Val Ile Val Ser Gln Ile His Arg Ser Pro
357       130           135           140

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/783,206

DATE: 11/17/2004
TIME: 12:20:53

Input Set : A:\Ipt01202.app
Output Set: N:\CRF4\11172004\J783206.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 1247
Seq#:2; Xaa Pos. 695,696
Seq#:5; Xaa Pos. 4
Seq#:6; Xaa Pos. 8,9,12,20,24,27,29,30,31,32
Seq#:7; Xaa Pos. 2,3,11,12
Seq#:8; Xaa Pos. 4
Seq#:13; Xaa Pos. 4
Seq#:14; Xaa Pos. 3,6,7,8,9,10,11,12,14,17,18,21,22,23,28,30

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/783,206

DATE: 11/17/2004

TIME: 12:20:53

Input Set : A:\Ipt01202.app

Output Set: N:\CRF4\11172004\J783206.raw

L:263 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:1232
L:461 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:688
L:649 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0
L:714 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0
M:341 Repeated in SeqNo=6
L:745 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0
L:760 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0
L:821 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
L:911 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0
M:341 Repeated in SeqNo=14